Video Structured Description System with Private Cloud

Qianjin Tang

R. &D. Center of Internet of Things Technology
The Third Institute of Ministry of Public Security
Shanghai, China
e-mail: Tangqj2008@163.com

Abstract—**Video surveillance system is more and more significant in public safety currently, but the video is unstructured data, which can not meet the demand of data mining applications. The text description of video content is a feasible direction. Cloud computing is one of technologies to solve the computing demand for video content analysis. The bandwidth is the main limitation for cloud computing of video surveillance system. A video structured description prototype system with private cloud was put forward. The resource scheduling module and video segmentation processes are introduced in detail. The result shows that private cloud is more suitable for video surveillance system.**

Keywords-private cloud; cloud computing; video surveillance; video structured description; VSD; video content analysis

I. INTRODUCTION

In recent years, video surveillance market has achieved rapid development with applications of by "Safety City", "3111 Project" and other state projects as well as the Olympic Games, World Expo in China. Police are turning to high-fidelity video surveillance technologies to monitor and prevent crime and profiling social behavior. Technology innovation is driving toward high-definition video surveillance, intelligent video surveillance, IP network direction, and video surveillance is moving from "visible" toward "see clearly", and even toward "able to understand" development. Video analysis technology is the basis of video understands. In fact, the automatic detection of surveillance related events from video input is an important technology in CCTV security applications, which has been proved to be feasible [1]. The techniques of machine vision [2] and images understandings [3, 4] are very significant to anti-terrorism and domestic safety.

However, there are still some difficulties and restrictions for the further applications of video surveillance system. The most prominent of limitations is that video is unstructured data, which is difficult to conduct data mining and difficult to be integrated with existing applications systems. For the private video network of the video surveillance system in China, the application requirement of the police is that the video surveillance system can be regarded as an information resource and integrated with other information resource.

The idea of video structured description (VSD) is to convert the contents of the video to text which can be data mining. VSD process has higher requirements for computing resources and cloud computing provides technical support with the reasonable computing resources scheduling. But cloud computing should take full account of the conditions including idle computing resources, bandwidth, and so on. For video surveillance system, the private cloud is more feasible for the bandwidth limitations.

II. VIDEO STRUCTURED DESCRIPTION

VSD was first proposed for visual surveillance system by the Third Research Institute of Ministry of Public Security in 2009 in China. VSD is a technology of putting video content into the text information which understood by the human and computer according to the semantic relation and adopt spatio-temporal segmentation, feature extraction, object recognition and so on[5,6]. VSD includes two meanings: one is the video semantic content, namely under the standardized video content description standard organization, each interested in the video of the target and its characteristics and behavior is identified, in the form of text to describe the content of the video, which is a video of information extraction process; Second, correlation of video resources, set up single semantic interconnection camera video resources, make use of data mining methods for effective analysis and semantic retrieval become possible, also makes the video resource semantic interconnection with other information systems resources possible, this is a video information organization, management and mining, and auxiliary process of business requirements. The basic idea of VSD is illustrated by the Figure 1. Before original video data is pushed to storage unit, they are inputted into a processor for analyzing. We call this processor the VSD device. Information of video content are extracted here and organized for storage with assistance of a certain semantic model. There are two outputs of the description unit: video, and its corresponding description text. The text channel is parallel to the transmission channel of the duplicated video. That is to say, the VSD bypasses the existing video analysis system and we can keep the system compatibility to the most extent. Correspondence between video and description data is also kept by the system for further applications.
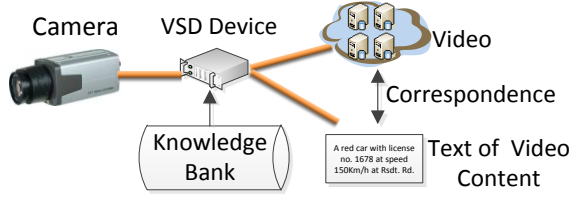
Figure 1. The schematic diagram of VSD technique

Different from other content-based intelligent video analysis techniques, by mining the prior field knowledge and scenario characteristic, according to the predefined rules for specific events, the VSD technique can use the standard metadata to describe the occurrence, the related people, vehicle, objects and the corresponding scenario overview. In others words, the VSD can translate the video information to structured text information composed by the elements related to the event itself. By utilizing the VSD technique, the surveillance video information can be extracted and processed automatically online and prompt warning message can be produced. We have built several VSD systems for practical applications in transportation management, jail management and service area surveillance in highway. The built VSD systems has greatly improved the management efficiency and enhanced the ability in information acquisition. The output text is shown by figure 2 which include information of licensee number, color, speed, time, day, direction, location, et al.

Figure 2. The output of a car in VSD system

III. CLOUD COMPUTING

Cloud computing is changing more and more services [7,8]. Cloud computing simply means that an application is running on a remote server instead of on a local computer, and the user accesses it via the Internet, much the way online banking is accomplished. In general, cloud computing is defined as using a network of remote servers hosted on the Internet to store, manage, and process data.

The advantages of video surveillance system with cloud computing over older technologies include:

- Flexible system size. Systems can be 1 camera at 1 location or 1,000 cameras at many locations, all managed through one interface online. This is opposed to DVR/NVR hardware which comes with limits as to the number of cameras and overall processing capability.
- Easy to expand. Simply add another camera and give it access to the internet and power. If you have an old DVR/NVR system that is maxed out, you can add a cloud camera to easily and inexpensively expand your number of cameras.
- Access your video from anywhere. Use any device that has access to the internet.
- Low cost. There's no need to buy expensive DVR/NVR hardware.
- IT friendly. There's no software to install for the basic system. Maintenance and upgrades are handled by the cloud server's owner.
- Easily Upgraded. Add functionality and convenience to your system by installing apps on your tablet or smart phone. Or install user viewing software on your PC. These upgrades do not interfere with browser access and are commonly free.

Cloud computing has many advantages with the assumption that network connectivity is going all times, but it is too ideal. In fact, the many conditions need to be considered when we apply cloud computing. Therefore, some believe the cloud computing applications is exaggerated in the video surveillance system [9]. An assessment was conducted to determine if it is practical for a video surveillance system to use various cloud technologies and answer the following two questions:

Is bandwidth enough？Cloud video surveillance system transfer "heavy" video data across different networks? Without enough bandwidth, you have no cloud video surveillance or, even worse, an unreliable solution. If you look at bandwidth as your "pipe" to the Internet, you need a really good "pipe" to get video up to the cloud. Your pipes have both upload and download speeds, which are usually different, and to get video to the cloud you need sufficient upload speed. The video stream of 720P HD video is over 4M. Therefore, for small-scale real-time transmission of the video contents, the current bandwidth can meet the demand of upload and download video, otherwise it is not.

Is resource idle? If there is no idle storage and computing resource, cloud computing is impossible.

The current video surveillance storage capacity requirements are set in accordance with the configuration, and generally without the remaining resources. But for video analysis computing resources, for the differences in analysis scenario of the video, there will be a number of idle computing resources. For example on the four lanes of same highway, each lane installs a surveillance camera, and each camera is equipped with computing resources device, but there may be at this moment, only two vehicles on the lane, then only two VSD device need start to compute, so that a surplus of the rest of the computing resources.

IV. PRIVATE CLOUD FOR VSD APPLICATION

A. System Architecture

The VSD system with private cloud has components including cameras, VSD devices, cloud management system, storage device, video application management system, fiber, and so on as shown in figure 3.
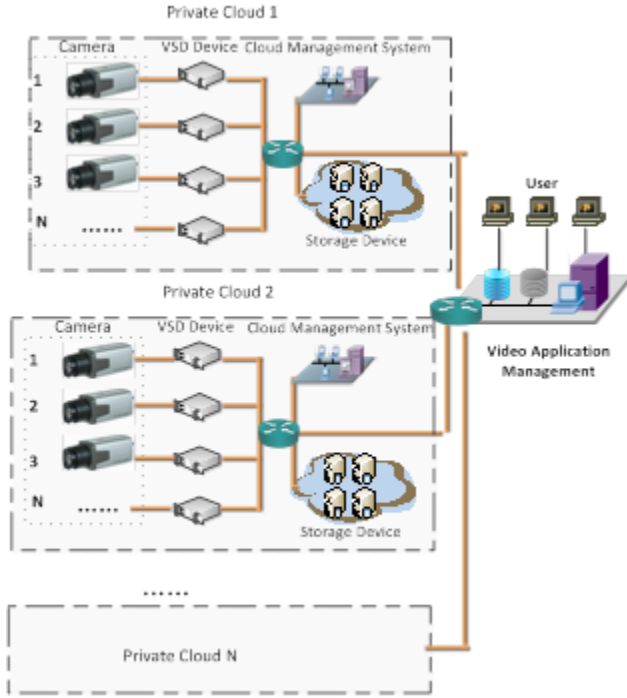
Figure 3: System architecture of private cloud for VSD application

The VSD system with private cloud is composed of two parts including private cloud and video application management platform. A private cloud section includes cameras, VSD device (i.e. video description computing resources), storage devices and cloud management systems and other accessories. Multiple private cloud application management systems share a video application management system which includes Web server, GIS server, user management server, security serves, and video streaming server.

A private cloud system are generally used in relatively close cameras, such as a road traffic junctions, and devices are connected through fiber interconnection, so the bandwidth will not be a problem, and the video resources can be real-time scheduling. Meanwhile, the same intersection as described in each camera scene at the same time, there may be busy with difference, so you can afford for the deployment of computing resources.

B. Resource Scheduling Management

The figure 4 illustrates the resource scheduling process. When there is a VSD task, the VSD device assess whether it need private cloud computing resources according to the computing complexity of the task. If the VSD device can meet the current real-time processing of video resources, the resource scheduling module (RSM) calculates the remaining capacity of computing power and saves it in computing capacity bank. If the device can not meet current resource VSD real-time processing, RSM starts the video segmentation module and a portion of the video is processed by the private cloud. You can also set the priority queue with a priority tag into the task queue. The queue priority

increases with time, so you can prevent the "hungry death "of the queue. Emergency task has a very high priority, and system you can direct request computing resources to complete the description and bypass the general task queue.
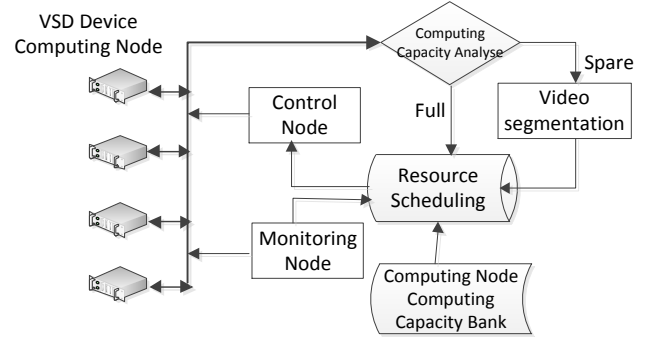

Figure 4. The technical architecture of resource scheduling

All computing nodes (VSD devices) upload their own situation to RSM by heartbeat model, including the current load, task priorities, and task execution schedule. RSM collects information through the monitoring node to obtain the cluster load distribution, and real-time distribution of cluster computing resources available according to computing power of the node information stored in the database. When there is a new task with particular high priority, the RSM can pause the cluster description for the lower priority task and serve the emergency task. Finally, a series of control commands are issued to the node control module.

Node control module is controlled by the RSM and controls compute nodes perform, pauses or resumes a particular task, including video pre-processing, video segmentation, video structured description processing. The control commands specify the tasks of the input and output, and the parameters are set in addition to basic operations such as perform, pause, and resume.

C. Video Segmentation

When the video file is split, it need judge the group structure of image stream. We should adopt different strategies for Open Group of Pictures (GOP) and Close GOP respectively. Open GOP need added one GOP redundancy at the top for each fragment to ensure that the decoder is capable of solving all the frames in the fragment. At the same time, system should record segmentation choice strategy for merging process reference. If the split point after the first GOP is Close GOP, it doesn't need choose other strategies. Conversely, if the split point after the first GOP is Open GOP, it need add a video after before a GOP.

V. CONCLUSION

The development of variety of video analysis technologies, such as a video surveillance system implanted in a strong intelligent core, make video surveillance truly real-time monitoring, real-time early warning capacity in the liberation of a large number of monitoring personnel, but

also has some difficulties. The VSD system of video content needs is more and more prominent. The high-definition video surveillance application brings clearer video sources, and is beneficial to structural description, while it has high demand for bandwidth. It is feasible for VSD system with private cloud.. This paper only put out a prototype system, which is currently under the project to optimize the design.

REFERENCES

[1] A. Mecocci, M. Pannozzo and A. Fumarola, "Automatic detection of anomalous behavioural events for advanced real-time video surveilance," International Symposium on Computational Intelligence for Measurement Systems and Applications, Lugano, pp. 187-192, July 2003.

[2] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, "Content-based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, pp.1349-1379, Dec. 2000.

[3] S. Zhu, D. Mumford, "A stochastic grammar of images," Foundations and Trends in Computer Graphics and Vision, vol.2, pp. 259–362, 2007.

[4] J. Sivic, B.C. Russell, A.A. Efros, A. Zisserman, W.T. Freeman, "Discovering object categories in image collections," Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp.734-741, Feb. 2005,.

[5] L. Mei, X. Cai, H. Zhang, C. Hu, and L. Qi, "Video structured description—vitalization techniques for the surveillance video data," W. Zhang et al. (Eds.): 9th International Forum on Digital TV and Wireless Multimedia Communication（IFTC 2012）, pp. 219–227, Nov. 2012.

[6] Z. Xu, Fenglin Zhi, C. Liang, and L. Mei, "Building video service platform in highway based on the video structural description technology," Journal of Network & Information Security, vol.2, pp. 64-72, 2011.

[7] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph,R. Katz, A. Konwinski, and M. Zaharia, "A view of cloud computing," Communications of the ACM., vol. 53, pp.50-58, 2010.

[8] I. Foster,Y. Zhao,I. Raicu, and S. Lu, "Cloud computing and grid computing 360-degree compared," Grid Computing Environments Workshop, vol. 901, pp.1-10, Nov 2008.

[9] D.J. Neal1, S. Rahman, "Video surveiliance in the cloud?" International Journal on Cryptography and Information Security (IJCIS), Vol. 2, pp.1-19, September 2012.